Logit Models for Bankruptcy Data Implemented in XploRe

A Master Thesis Presented

by

TALEB AHMAD
(185159)

to

Prof. Dr. Wolfgang Härdle

CASE - Center for Applied Statistic and Economics
Institute for Statistics and Econometrics

in partial fulfillment of the requirements

for the degree of

Master of Science

Humboldt-Universität zu Berlin

Berlin, 24th November 2005

Declaration of Authorship

I hereby confirm that I have authored this master thesis independently and without use of others than the indicated resources. All passages which are literally or in general matter taken out of publications or other resources are marked as such.

TALEB AHMAD

Berlin, 24th November 2005

Abstract

Numerous reasearch attempts in predicting business failures and or bankruptcy are well documented in corporate finance. Attempts to develope bankruptcy prediction continues since commercial banks, public accounting firms, bond rating agencies, for example have advocated for such information to minimize their exposure to potential client failures. The evolution of bankruptcy prediction research is geared towards the types of models that include statistical models (primarily, multiple discriminant analysis [MDA], conditional logit regression analysis, artificial neural network models and support vector machines [SVM]. Many additional bankruptcy model have been the work of Platt & Platt (1980), Gilbert, Menon, and Scwhartz (1990). Almost universally, the decision criteria to evaluate the usefulness of these models has been how well they classify a company as bankrupt or non-bankrupt compared to the company's actual status known after the fact. In this thesis I employ logit analysis as an easily implemented analytical procedure to a bankruptcy data with the use of the XploRe software. The content is as follows: Chapter 1 and 2 introduce some models and methods used to analyse binary data and describe some stochastic properties of these models. Chapter 3 introduces data preparation for the Bankruptcy data set used in this work. In chapter 4, I examines some binary model applications in XploRe. Chapter 5 presents the logit model estimation for the Bankruptcy data. Some cases of link function and Conclusion of the analysis is in chapter 6.

Keywords: logit model, probit model, bankruptcy

Acknowledgements

I would like to gratefully acknowledge the help of Prof. Dr. Wolfgang Härdle for the enthusiastic supervision, orientation offered and penetrating criticism. I thank to all my friends in the Institute of Statistics and Econometrics who have assisted me in the course of this study. Finally, this thesis is dedicated to my family, especially my father in Lattakia, Syria.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In past years, analysts relied principally on financial statements to evaluate risks associated with investment. For example, simple ratio analysis was performed to consider if the company was sufficiently liquid and to see how well it managed its assets and debt. It has been observed that ratio analysis is fairly meaningless taken alone. More recently, logit analysis has been compared to more advanced analytical tools, neural networks, support vector machines. Research has found that the approaches perform similarly well (see, Altman, Marco, and Varetto 1994, 505). Logit analysis actually provides a probability (in terms of a percentage) of bankruptcy. Also, the probability calculated might be considered a measure of the effectiveness of management, effective management will not lead a company to the verge of bankruptcy. This thesis considers the logit model approach to analyze a bankruptcy data. In addition we verify our results with the probit and a generalized partial linear model. A more advanced company classification method with support vector machines is also reflected in this work.

1.1 Bankruptcy Prediction Models

Attempts to develop bankruptcy prediction models began in the late 1960's and continue through today, most of the publicly available information regarding prediction models is based on research published by university professors. Commercial banks, public accounting firms. There are two main approaches in bankruptcy prediction studies : The first and most often used approach has been the empirical search for predictors (financial ratios) that lead to lowest misclassification rates. The second approach has concentrated on the search for statistical methods that would also lead to improved prediction accuracy. Bankruptcy prediction models are more generally known as measures of financial distress. Three stages in the development of financial distress measures exist: univariate analysis, multivariate analysis, and logit analysis. Univariate analysis assumes "that a single variable can be used for predictive purposes" (Cook and Nelson 1998). The univariate model as proposed by William Beaver achieved a "moderate level of predictive accuracy" (Sheppard 1994, 9). By this framework Beaver state four propositions:

1. The larger the reservoir, the smaller the probability of failure.
2. The larger the net liquid-asset flow from operations, the smaller the probability of failure.
3. The larger the amount of debt held, the greater the probability of failure.
4. The larger the fund expenditures for operations, the greater the probability of failure.

Beaver identified 30 ratios that were expected to capture relevant aspects. By a univariate discriminant analysis, these ratios were applied on 79 pairs

of bankrupt/nonbankrupt firms. The best discriminators were working capital funds flow/total assets and net income/total assets which correctly identified 90% and 88% of the cases. Univariate analysis identified factors related to financial distress; however, it did not provide a measure of the relevant risk (Stickney 1996, 507). The studies discussed before make use of several different ratios. These ratios tell us something about the probability of bankruptcy. Most of these ratios measure profitability, liquidity, and solvency. The aforementioned studies did not make clear which ratios have the most explaining power. For that reason, we have the next question: which ratios are most important in the prediction of bankruptcy. In the next stage of financial distress measurement, multivariate analysis (also known as multiple discriminant analysis or MDA) attempted to "overcome the potentially conflicting indications that may result from using single variables" (Cook and Nelson 1998). Multiple discriminant analysis method is the one proposed by Edward Altman. Altman's z-score, or zeta model, combined various measures of profitability or risk . The resulting model was one that demonstrated a company's risk of bankruptcy relative to a standard. Altman was using the 7 ratios; return on assets, stability of earnings, debt service, cumulative profitability, liquidity, capitalization and size. Applied on 33 pairs of bankrupt/non-bankrupt firms the model correctly identifies 90% of the cases one year prior to failure. Although the positive results of his study, Altman's model had a key weakness: it assumed variables in the sample data to be normally distributed. "If all variables are not normally distributed, the methods employed may result in selection of an inappropriate set of predictors" (Sheppard 1994). Ohlson(1980) is the first to apply the logit analysis on the problem of bankruptcy prediction. By using 105 bankrupt and $2,058$ non-bankrupt firms he is also the first to apply a representative sample. He states that predictive power appears to be less than reported in previous studies. Further, logit analysis actually provides a probability (in terms of a percentage) of bankruptcy, the probability calculated might be considered a measure of the effective-

ness of management. During the $1980s$ and $1990s$, the trend has been to use logit analysis in favor of multiple discriminant analysis (Stickney 1996, 510). Logit analysis has been compared to a more advanced analytical tool, neural networks. Research has found that the approaches perform similarly and should be used in combination (Altman, Marco, and Varetto 1994, 505).

1.2 The Bankruptcy Laws and the Financial Regulation

Operating firm must be solvent. According to accounting principles it signifies that it can serve and refund all its debt's becoming due. Insolvency is the situation when the firm's debt is greater than its asset value including: stocks, accorded credits, real estates, machines and other assets. In a situation of insolvency bankruptcy occurs. It is the pattern for resolving disbursement problems of firm owners. Historically, bankruptcy consists in three stages:

1. To withdraw publicly the bankrupt from operating.
2. To gather all information about creditors and to estimate assets.
3. To settle the investors' failure (they lose for this reason their property rights), to sale assets in order to indemnify the creditors, to quash the marginal debts and to arrange the firm's liquidation. (See Peaucelle, 2005).

Two principal forms of bankruptcy procedures exist: an asset sale and a structural bargaining. The sale of the firm's assets is usually supervised by a trustee, or a receiver. Such procedures and supervision are not as evident as it seems from the old capitalist world. The bankruptcy reforms are

in progress in many countries in order to make the procedures more transparent and efficient (see Hart, 1999). Thus, the goal of an appropriate bankruptcy law is to reduce the systemic risk and overall financial instability. But the regulation of financial system, with the prudential rules is another way. The widespread financial distress may come from the failure of individual institutions and the spread trough different contagion mechanisms to the financial system in general (Gourieroux & Peaucelle (1996)).

Chapter 2

Methodology

2.1 The Logit and Probit models

There is an alternative interpretation that gives rise to the probit model. Consider a latent variable

$$y_i^* = x_i^\top \beta + \varepsilon_i$$

That linearly depends on x_i and the error term $\varepsilon_i \sim N(0, \sigma^2)$. Choosing the case $y_i = 1$ if the latent variable is positive and 0 otherwise, we have the form

$$y_i = \begin{cases} 1 & y_i^* > 0 \\ 0 & y_i^* < 0 \end{cases}$$

The latent variable interpreted as the utility difference between choosing $y_i = 1$ and 0. The probability that $y_i = 1$ can be derived from the latent variable and the decision rule.

$$\begin{aligned} P(y_i = 1 \mid x_i) &= P(y_i^* > 0 \mid x_i) \\ &= P(x_i^\top \beta + \varepsilon_i > 0 \mid x_i) \\ &= P(\varepsilon_i > -x_i^\top \beta \mid x_i) \\ &= 1 - \Phi\left(-\frac{x_i^\top \beta}{\sigma}\right) \\ &= \Phi\left(\frac{x_i^\top \beta}{\sigma}\right) \end{aligned}$$

Assuming that the error term has a standard normal distribution $\varepsilon_i \sim N(0,1)$, we have the equation

$$\pi_i = \Phi\left(\eta_i\right)$$

Where Φ is the standard normal $c.d.f$. The inverse transformation which gives the linear predictor as a function of the probability is

$$\eta_i = \Phi^{-1}(\pi_i)$$

The transformation function in the probit model is the cdf of the standard normal distribution

$$\begin{aligned} P(y_i = 1 \mid x_i) &= \Phi\left(x_i^\top \beta\right) \\ &= \int_{-\infty}^{x_i^\top \beta} \Phi(z) d_z \end{aligned}$$

An alternative model is the logit model that uses the logistic function

$$\begin{aligned} P(y_i = 1 \mid x_i) &= G(x_i^\top \beta) \\ &= \frac{e^{x_i^\top \beta}}{1 + e^{x_i^\top \beta}} \\ &= \frac{1}{1 + e^{-x_i^\top \beta}} \end{aligned}$$

If the error term has a standard normal distribution, we have the probit model, and if the error term has a logistic distribution, we have the logit model.

2.2 Bernoulli and Binomial Distribution

For a randomly-selected individual from the population of bankruptcy data, Y is a binary $(0/1)$ random variable, in the population Y that can take the values one and zero with probabilities p_i and $1-p_i$, the distribution of Y_i is called a Bernoulli distribution with parameter p_i with

$$P(Y_i = y_i) = p_i^{y_i}(1-p_i)^{1-y_i} \tag{2.1}$$

For $y_i = 0, 1$. Note that if $y_i = 1$ we obtain p_i, and if $y_i = 0$ we obtain $1-p_i$, on the other hand randomly-draw a sample of n individuals from the population where $P(Y = 1) = p_i$. Let n binary results be $Y_1, Y_2, ..., Y_n$. Now if the n individuals are independent and if the n individuals all have the same probability of bankruptcy (probability that $Y_i = 1$), then Y has a Binomial distribution with parameters p_i and n_i, which we can write $Y_i \sim \beta(n_i, p_i)$. The probability distribution function of Y_i is given by

$$P(Y_i = y_i) = (y_i^{n_i})p_i^{y_i}(1-p_i)^{n_i-y_i} \tag{2.2}$$

The mean and variance of Y_i can be shown to be

$$E(Y_i) = \mu_i = n_i p_i \tag{2.3}$$

$$Var(Y_i) = \sigma_t^2 = n_i p_i (1-p_i) \tag{2.4}$$

Respectively. For our data one can write the form of Bernoulli distribution as

$$y_i = \begin{cases} 1 & \text{if a company went bankruptcy within three years} \\ 0 & \text{if it survived} \end{cases}$$

2.3 The Logit Transformation

Logistic regression is a technique for analyzing problems in which there are one or more independent variables that determine an outcome. The outcome is measured with a dichotomous variable (in which there are only two possible outcomes). The goal of logistic regression is to find the best fitting model to describe the relationship between the dichotomous characteristic of dependent variable and a set of independent variables. Logistic regression generates the coefficients, its standard errors and significance levels of a formula to predict a logit transformation of the probability of presence of the characteristic of interest. Now we have the following equation that p_i be a linear function of the covariates.

$$p_i = X_i^\top \beta \tag{2.5}$$

where β is a vector of regression coefficients. The equation (2.5) is called the linear probability model. This model can estimated from individual data using ordinary least squares (OLS). We have one problem with this model is that the probability p_i on the left-handside has to be between zero and one, but the linear predictor $X_i^\top \beta$ on the right-hand-side can take any real value. Thus there is a simple solution to this problem that one can to transform the probability to remove the range restrictions, and model the transformation as a linear function of the covariates. We do this in two steps. First, we move from the probability p_i to the odds, $odds_i = \frac{p_i}{1-p_i}$ defined as the ratio of the probability to its complement, second, we take logarithms, calculating the logit or log-odds

$$\eta_i = p_i = \log \frac{p_i}{1 - p_i} \tag{2.6}$$

We can note that the probability goes down to zero the odds approach zero and the logit approaches $-\infty$. At the other extreme, as the probability approaches one the odds approach $+\infty$ and so does the logit. Thus, logits

map probabilities from the range $(0, 1)$ to the entire real line. Negative logits represent probabilities below one half and positive logits correspond to probabilities above one half. Figure (2.1) illustrates the logit transformation. In the bankruptcy use data there are 426 companies went bankrupt among 1052 company, so we estimate the probability as $426/1052 = 0.41$. The odds are $426/626$ or 0.68 to one, the logit is $\log(0.68) = 0.38$. The logit transformation is one-to-one. The inverse transformation is sometimes called the antilogit, and allows us to go back from logits to probability.

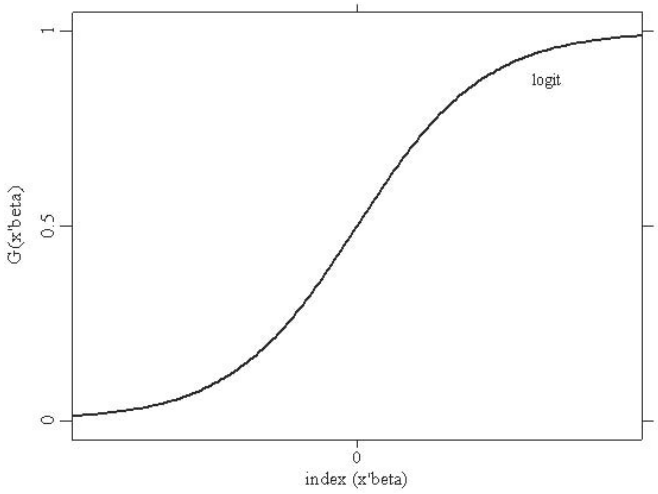

Figure 2.1: The logit transformation

logittransformation.xpl

$$p_i = Logit^{-1}(\eta_i) = \frac{e^{\eta_i}}{1 + e^{\eta_i}} \tag{2.7}$$

In our data the estimated logit was 0.38. Exponentiating this value we obtain the odds and a probability 0.41 .

2.4 The Logistic Regression Model

The logistic regression model describes the relationship between a dichotomous response variable Y, coded to take the values 1 or 0 for success and failure, and K explanatory variables $x_1, x_2, ...x_k$. The explanatory variables can be quantitative or indicator variables referring to the levels of categorical variables. Since Y is a binary variable, it has a Bernoulli distribution with parameter $p = P(Y = 1)$, that is, p is the probability of bankruptcy for given values $x_1, x_2, ...x_k$ of the explanatory variables. Suppose that $Y_1, ..Y_n$ are independent Bernoulli variables. We can say that Y_i has a Binomial distribution

$$Y_i \sim \beta(n_i, p_i) \tag{2.8}$$

We can suppose that the logit of the probability p_i is a linear function of the predictors

$$Logit(p_i) = X_i^\top \beta \tag{2.9}$$

Where x_i is a vector of covariates and β is a vector of regression coefficients. The model defined in equations (2.8) and (2.9) is a generalized linear model with binomial response and link logit. β_j represents the change in the logit of the probability associated with a unit change in the j-th predictor holding all other predictors constant. Exponentiating equation (2.9) we find that the odds for the i-th unit are given by

$$\frac{p_i}{1 - p_i} = e^{X_i^\top \beta}$$

Last equation defines a multiplicative model for the odds. When one can solving for the probability p_i in the logit model in equation (2.9) gives the more complicated model

$$p_i = \frac{e^{X_i^\top \beta}}{1 + e^{X_i^\top \beta}} \tag{2.10}$$

we can see the left-hand-side is in the familiar probability scale, and the right-hand side is a non-linear function of the predictors, and there is no simple way to express the effect on the probability of increasing a predictor by one unit while holding the other variables constant.

2.5 Maximum Likelihood Estimation

We can define the likelihood function by the next form

$$\begin{aligned} L(x \mid \theta) &= P(x \mid \theta) \\ &= P(x_i \mid \theta) \end{aligned}$$

We call $L(x \mid \theta)$ is the probability that the data x is observed, given that the parameter value is θ. The maximum likelihood estimator (MLE) is derived by holding x fixed and maximising L over all possible values of θ

$$\theta_{MLE}(x) = \arg\max L(x \mid \theta)$$

The maximum likelihood estimate is the value of θ for which the associated distribution (among all distributions parameterised by θ) is most likely to have generated the data x. We can consider the family of Binomial distributions as follows

$$P = \beta(n, \theta) : \theta \in [0, 1]$$

Where n is the number of trials and θ is probability of bankruptcy. The likelihood function is

$$L(y_i \mid \theta) = (y_i^{n_i})\theta^{y_i}(1-\theta)^{n_i - y_i} \tag{2.11}$$

Since $\log(L)$ is a monotonic increasing function of L, the value of θ that maximises L also maximises $\log(L)$. To find this value, we differentiate $\log(L)$

$$\begin{aligned} \frac{\partial}{\partial\theta}\log(L) &= \frac{\partial}{\partial\theta}\left\{\log(y_i^n) + y_i\log\theta + (n-y_i)\log(1-\theta)\right\} \\ &= \frac{y_i}{\theta} - \frac{n-y_i}{1-\theta} \end{aligned}$$

setting this equal to zero, we obtain the maximum likelihood estimate

$$\widehat{\theta}_{MLE} = \frac{y_i}{n} \tag{2.12}$$

the MLE for θ is therefore equal to the number of failure expressed as a proportion of the total number of trials.

2.6 Tests of Hypotheses

In logistic regression, hypotheses on significance of explanatory variables cannot be tested in quite the same way as in linear regression. Recall that in linear regression, where the response variables are normally distributed, we can use t- or F -test statistics for testing significance of explanatory variables. But in logistic regression where the response variables are Bernoulli distributed. We have to use different test statistics which exact distributions are unknown. One can use two different types of test statistics: The log likelihood ratio statistic and the Wald statistic. We can say that the likelihood statistic is superior to the Wald statistic because that it gives more reliable results, so we shall mainly concentrate on the likelihood ratio statistic. The reason for considering the Wald statistic too is that it is computationally easy and is given automatically in the output of most statistical computer packages. We can test the hypothesis

$$H_0 : \beta_j = 0$$

Concerning the significance of a single coefficient by calculating the ratio of the estimate to its standard error

$$Z = \frac{\widehat{\beta}_j}{\sqrt{Var(\widehat{\beta}_j)}}$$

This statistic has approximately a standard normal distribution in large samples. Alternatively, we can treat the square of this statistic as approximately a Chi-squared with one d.f. The Wald test can be use to calculate a confidence interval for β_j

$$\widehat{\beta}_j + Z_1 - \alpha/2\sqrt{Var(\widehat{\beta}_j)}$$

Where $Z_1 - \alpha/2$ is the normal critical value for a two-sided test of size α. Confidence intervals for effects in the logit scale can be translated into confidence intervals for odds ratios by exponentiating the boundaries.

Chapter 3

Data Preparation, the Bankruptcy Data

The bankruptcy data shows profitability and liquidity financial ratios of US successful and failing companies. In our analysis we consider the current state of bankruptcy as the response or dependent variable of 14 variables as predictors. The source for this data is annual reports of the companies from $1990-2004$ available from Compustat. In this data (Table 3.1) we have $n = 1052$ companies, around 426 companies with capitalization exceeding 1 billion went bankrupt in three years and there are 626 surviving companies of a similar size and the same industry according to the standard industrial classification code (SIC). Table 3.1 presents the description for these variables. Note that the response has two categories 1 if a company seek protection under chapter 11 of the US Bankruptcy code within three years, 0 otherwise. In our data the companies were characterized by 14 variables from which the following financial ratios as shown in table 3.1 were calculated:

Variable	Symbol	The description of the variables
X_1	Cash-TA	Cash/Total Assets
X_2	Inv-TA	Inventories/Total Assets
X_3	CA-TA	Current Assets/Total Assets
X_4	Kap-TA	Property, Plant and Equipment/Total Assets
X_5	Intg-TA	Intangibles/Total Assets
X_6	Log TA	log Total Assets
X_7	Cl-TA	Current Liabilities/Total Assets
X_8	TL-TA	Total Liabilities/Total Assets
X_9	Eq-TA	Equity/Total Assets
X_{10}	S-TA	Sales/Total Assets
X_{11}	Ebit-TA	EBIT/Total Assets
X_{12}	Ebit-Int	EBIT/Interest Payments
X_{13}	NI-TA	Net Income/Total Assets
X_{14}	CA-CL-TA	(Current Assets - Current Liabilities)/Total Assets
X_{15}	BANKR	Bankruptcy (1=bankrupt, 0=operating)

Table 3.1: Description the variables of the bankruptcy data $n = 1052$

1. Profit measures: the variables are X_{11} $(EBIT - TA)$, X_{13} $(NI - TA)$.

2. Leverage ratios: the variables are X_4 $(Kap - TA)$, X_8 $(TL - TA)$, X_9 $(Eq - TA)$.

3. Liquidity ratios: the variables are X_1 $(Cash - TA)$, X_3 $(CA - TA)$, X_7 $(CL - TA)$, X_{14} $(CA - CL - TA)$.

4. Activity or turnover ratios: the variables are X_2 $(Inv - TA)$, X_{10} $(S - TA)$, X_{12} $(EBIT - Int)$.

I have two questions in this study, the first question: which financial ratios have a big influence on the probability of bankruptcy . And the second question: which method is better for the classification problem for the bankruptcy data . I will try to answer these questions later.

Variable	Mean	Median	Variance	Sgrt	Skewness	Kurtosis
X_1	0.148	0.07	0.042	0.205	2.343	8.451
X_2	0.145	0.09	0.029	0.169	1.467	4.889
X_3	0.467	0.45	0.069	0.262	0.202	2.024
X_4	0.357	0.30	0.068	0.261	0.535	2.228
X_5	0.078	0	0.022	0.148	2.707	11.21
X_6	4.729	5.03	7.818	2.796	-0.318	2.635
X_7	0.406	0.24	0.502	0.709	6.686	60.15
X_8	0.965	0.69	1.135	1.065	4.287	28.10
X_9	0.035	0.31	1.135	1.065	-4.287	28.10
X_{10}	1.182	1.01	0.842	0.917	1.479	6.549
X_{11}	-0.145	0.02	0.878	0.937	-10.97	163.5
X_{12}	-593.1	0.55	6.622	8.137	-0.763	13.33
X_{13}	-0.188	0.01	7.819	2.796	-14.50	365.9
X_{14}	0.060	0.14	0.541	0.736	-5.942	52.69

Table 3.2: Descriptive statistics for bankruptcy data $n = 1052$

Chapter 4

Some Applications in XploRe

4.1 Bivariate Plots

Figure 4.1 shows a Bivariate plot for two variables, $(NI - TA)$ and $(TL - TA)$. The blue dots shows surviving companies where as the red dots shows the bankrupt companies within three years.

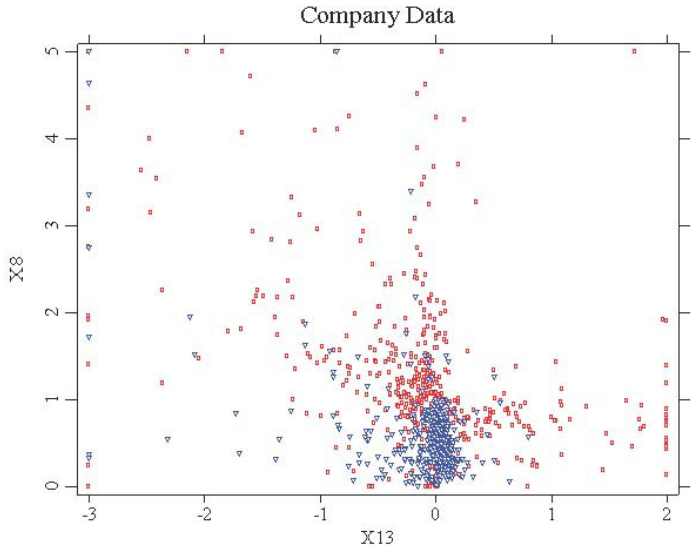

Figure 4.1: Bivariate Plot of the bankruptcy data

plotData.xpl

4.2 Scatter-Plot Matrices

Choosing 4 variables ($Cash - TA$, $Inv - TA$, $TL - TA$, $NI - TA$), we present a scatter plot, in figure 4.2 for every possible variable combination. With every variable there are two sorts of points: the red points indicate that a company went bankruptcy within three years, and the blue points it means if the company survived.

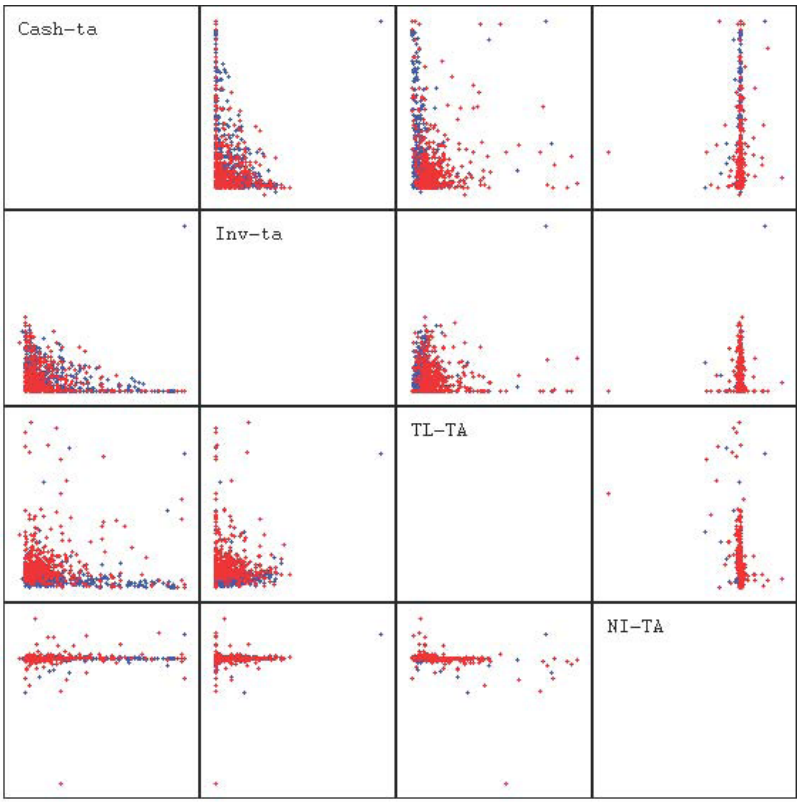

Figure 4.2: Scatter-plot matrix of the X_1 (Cash-TA), X_2 $(Inv - TA)$, and X_8 $(TL - TA)$, and X_{13} $(NI - TA)$ variables of the bankruptcy data

.

scattplot-bank2.xpl

4.3 Boxplot

Figure 4.3 shows boxplots for all variables of the bankruptcy data. From

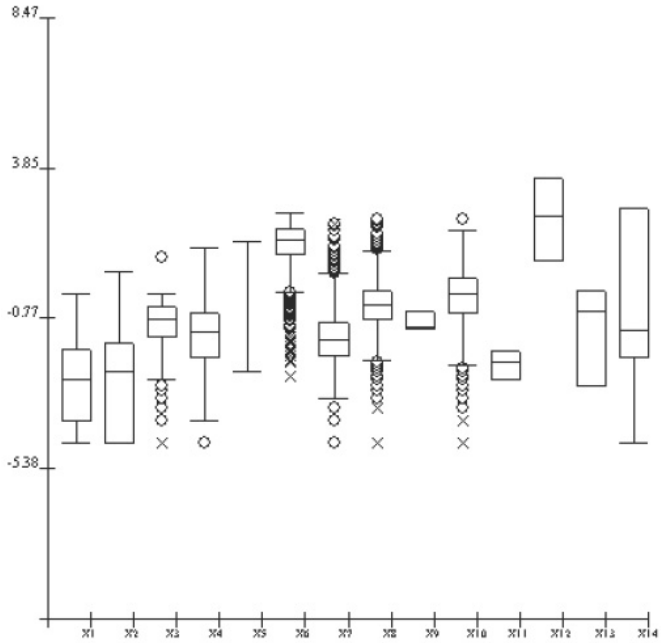

Figure 4.3: Boxplot for all variables of the bankruptcy data

boxplot1.xpl

this plot, we see that the variables X_6, X_7, X_8, X_{10} have some outliers. These outliers are marked with circles and crosses. On the other hand we can consider that variables $X_4, X_6, X_7, X_8, X_{10}$ have a symmetrical distribution because they have same distance from the median (solid line in these boxes). We make another figures for these variables, one can do two boxplots together for every variable, for example the variable X_8 (total liabilities to total assets ratio) as in the figure (4.4), the blue boxplot on the left indicate that the company was not bankruptcy, and the red boxplot on the right indicate that the company went bankruptcy, (the rule is similar in another variables X_1, X_2, X_3, X_4, X_6, X_7, X_{10}).

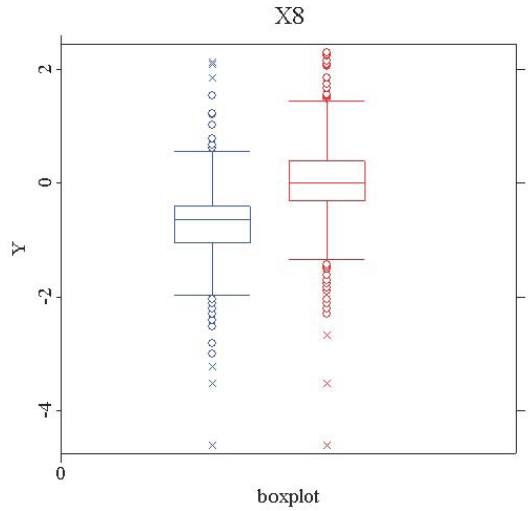

Figure 4.4: Boxplot for the variable X_8 of the bankruptcy data

boxplot2.xpl

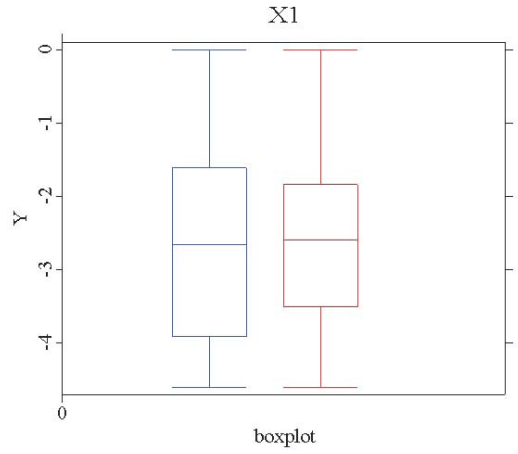

Figure 4.5: Boxplot for the variable X_1 of the bankruptcy data

boxplot2.xpl

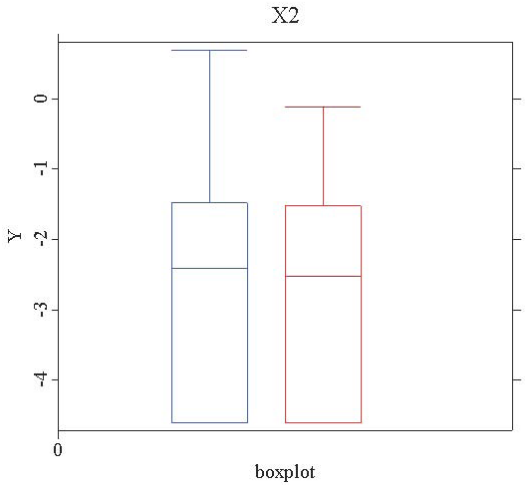

Figure 4.6: Boxplot for the variable X_2 of the bankruptcy data

boxplot2.xpl

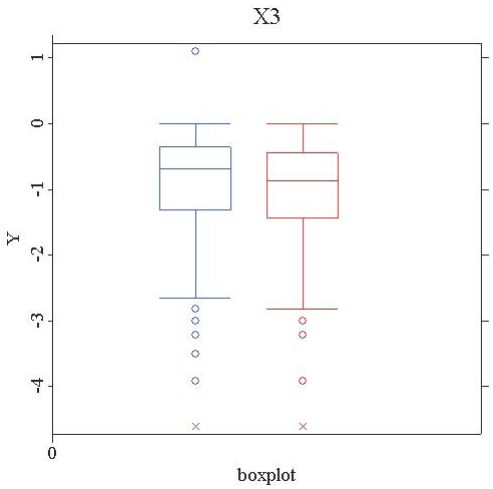

Figure 4.7: Boxplot for the variable X_3 of the bankruptcy data

boxplot2.xpl

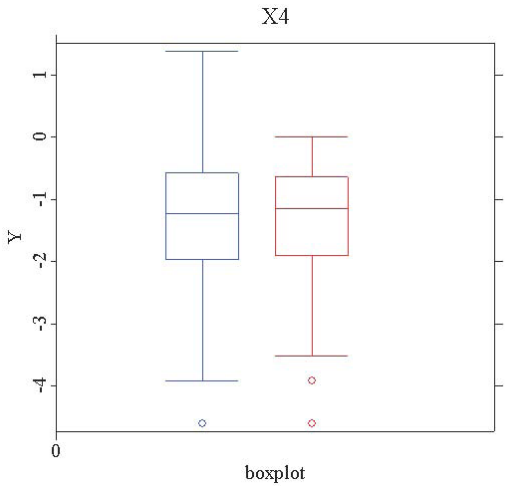

Figure 4.8: Boxplot for the variable X_4 of the bankruptcy data

boxplot2.xpl

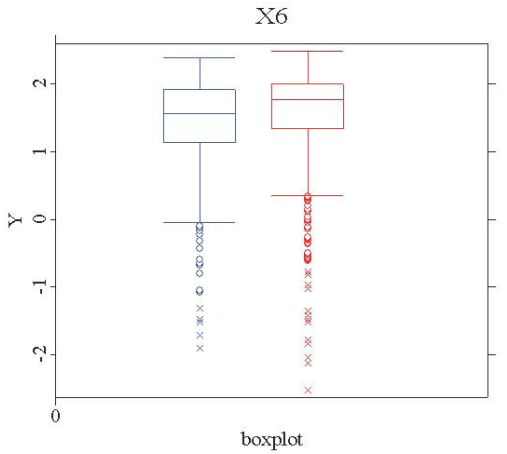

Figure 4.9: Boxplot for the variable X_6 of the bankruptcy data

boxplot2.xpl

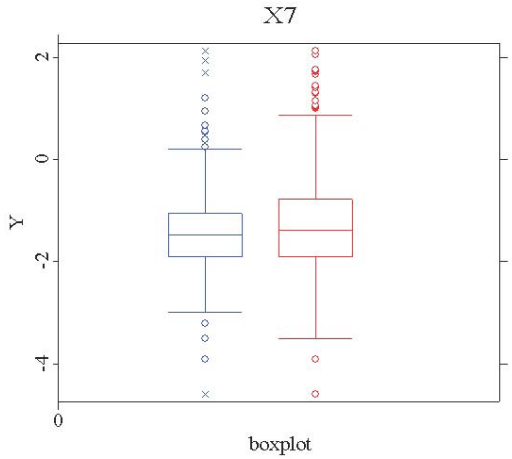

Figure 4.10: Boxplot for the variable X_7 of the bankruptcy data

boxplot2.xpl

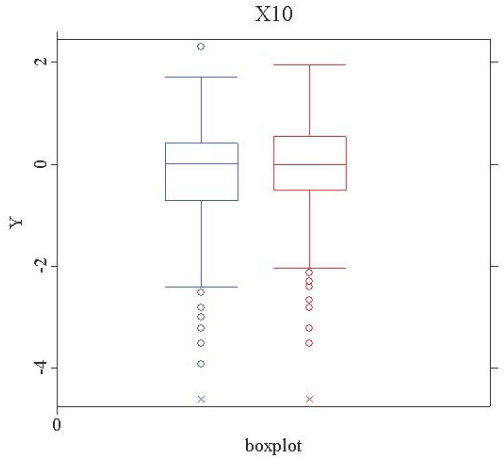

Figure 4.11: Boxplot for the variable X_{10} of the bankruptcy data

boxplot2.xpl

4.4 Company Classification with SVMs

Since their introduction in 1992, Support Vector Machines marked the beginning of a new era in the learning from examples paradigm in artificial intelligence. Rooted in the Statistical Learning Theory developed by Vladimir Vapnik, Support Vector Machines quickly gained attention from the pattern recognition community due to a number of theoretical and computational merits. Support Vector Machines represent a breakthrough in the theory of learning systems. Statistical Learning Theory, the backbone of Support Vector Machines, provides a new framework for modeling learning algorithms, merges the fields of machine learning and statistics, and inspires algorithms that overcome all of the above difficulties. A new generation of learning algorithms - or equivalently of statistical methods - has recently been developed, based on this theory. Such methods prove remarkably resistant to the problems imposed by noisy data and high dimensionality. They are computationally efficient. The optimal solution can always be found. These methods have an inherent modular design that simplifies their implementation and analysis and allows the insertion of domain knowledge. More importantly, they come with theoretical guarantees about their generalization ability. SVMs are a group of methods for classification (and regression) that make use of classifiers providing "high margin". SVMs possess a flexible structure which is not chosen a priori. To show the ability of an SVM to extract information from the data, we take tow ratios: $(NI - TA)$, $(TL - TA)$. Triangles in these figures represent successful companies and squares represent failing companies, the darker the area is the higher probability of bankruptcy. We see that the successful companies lying in the bright area have positive profitability, in these figures we see the effects of different classifier functions complexities according the radial basis is 100 in the figure (4.12) and 2 in the figure (4.13) and 0.5 in the figure (4.14) and the capacity is fixed $C = 1$. We see if the complexity of classifying

functions increases we get a better picture, and the areas of successful and failing companies become localized. We can work company classification

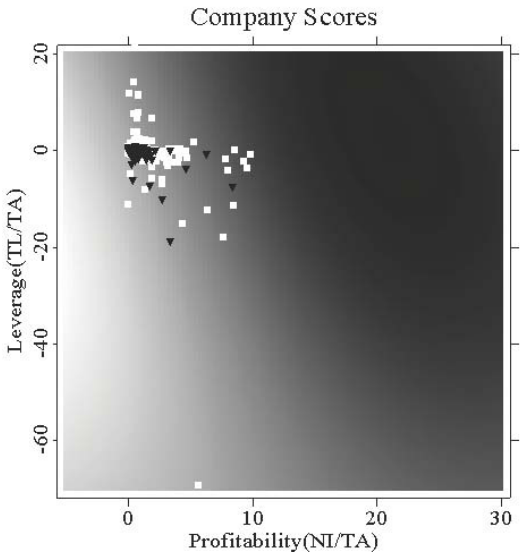

Figure 4.12: The case of a low complexity of classifier functions, the radial basis is 100 and $C = 1$

talebsvm.xpl

based on the effects of high capacity we choose $c = 300$ and the radial is 2, we get one cluster of successful companies and the cluster for bankrupt companies disappear. As the figure (4.15).

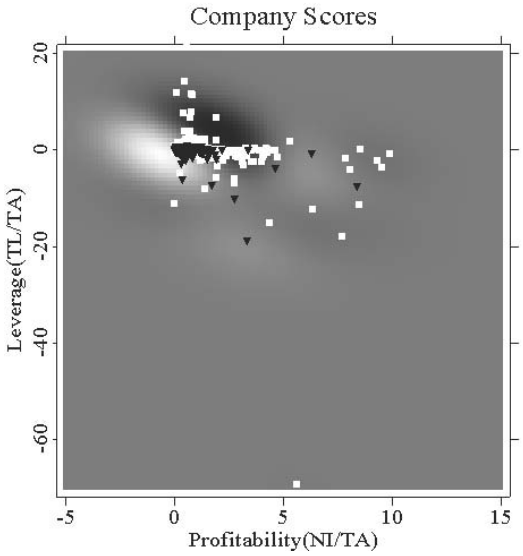

Figure 4.13: The case of an average complexity of classifier functions, the radial basis is 2 and $C = 1$

talebsvm.xpl

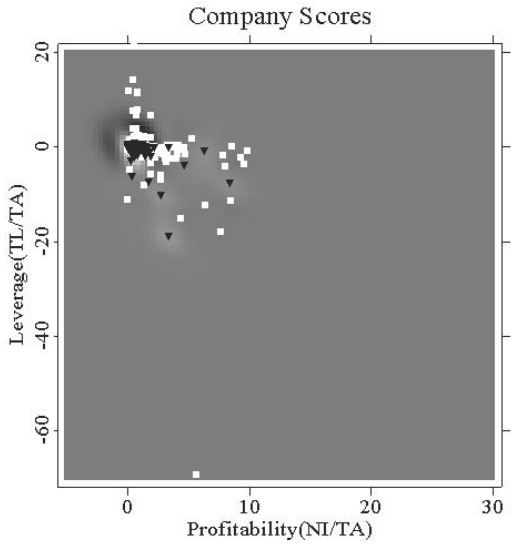

Figure 4.14: The case of excessively complexity of classifier functions, the radial basis is 0.5 and $C = 1$

talebsvm.xpl

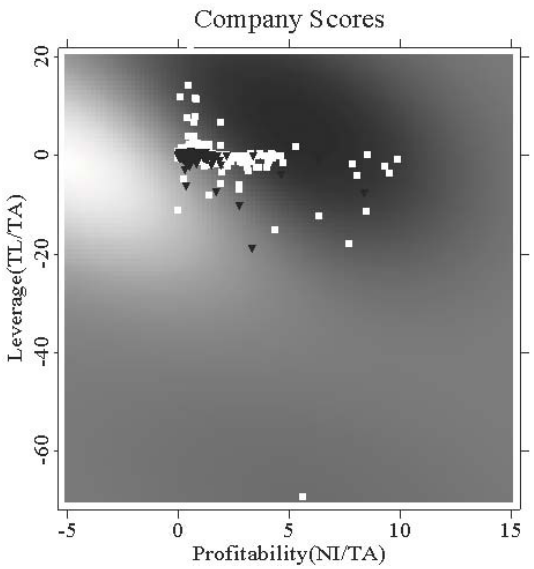

Figure 4.15: The case of high capacity, the radial basis is 2 and $C = 300$

talebsvm.xpl

Chapter 5

Computing GLM Estimates

5.1 Estimation Logit Model

Logit analysis has also been used to investigate the relationship between binary or ordinal response probability and explanatory variables. The method fits linear logistic regression model for binary or ordinal response data by the method of maximum likelihood. Among the first users of logit analysis in the context of financial distress was Ohlson (1980). Like discriminant analysis, this technique weights the independent variables and assigns a Z score in a form of failure probability to each company in a sample. The advantage of this method is that it does not assume multivariate normality and equal covariance matrices as discriminant analysis does. Logit analysis incorporates non-linear effects, and uses the logistical cumulative function in predicting a bankruptcy. For our data we estimate the logit model and present graphical output display in figure (5.1). This plot shows $X\beta$ vs the predicted regression function (green line). We can see in this plot a graphical representation of the bankruptcy. Each company is represented be a " $+$ ". Each company is transformed into an index laid on the horizontal axis and the dependent variable Y laid on the vertical axis. The output display shows the estimation results. The table (5.1) gives the es-

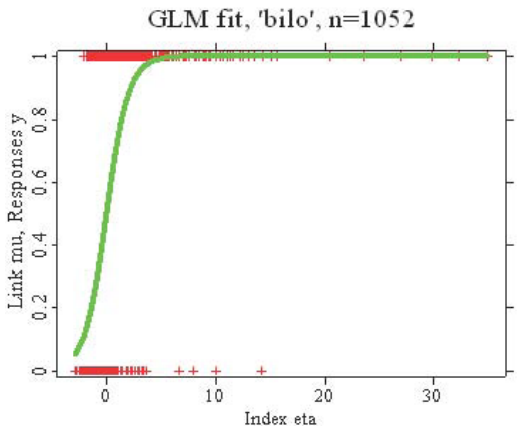

Figure 5.1: Logit fit

glm.xpl

timated coefficient vector β together with the estimated standard errors and t-values. The table (5.1) shows the results of this logit fit. We see that the variables X_8 (total liabilities to total assets ratios) has an important influence on bankruptcy state of company, another variables X_1 (cash to total assets ratio), X_2 (inventories to total assets ratio), X_{13} (net income to total assets ratio), X_4 (property, plant and equipment to total assets ratio), X_6 $(Log - TA)$ have strong influence, and one can see that variables are highly significant, which is indicated by their high t -values. Another variables have not enough influence. At the end we say that profitability, leverage and Liquidity have important effects on the probability of bankruptcy. Table (5.2) give some statistics for this fit, where R^2 is (the coefficient of determination), χ^2 is person statistic, and σ^2 is the variance.

Estimates	B	s.e.	t-value
[const]	-1.01	0.54	-1.87
[Cash-TA]	**3.17**	**0.73**	**4.33**
[Inv-TA]	**2.69**	**0.80**	**3.35**
[CA-TA]	-22.9	15.7	-1.46
[Kap-TA]	**-1.26**	**0.57**	**-2.20**
[Intg-TA]	0.38	0.71	0.53
[Log-TA]	**-0.10**	**0.03**	**-2.98**
[CL-TA]	17.9	15.6	1.15
[TL-TA]	**3.82**	**0.31**	**12.3**
[S-TA]	0.17	0.12	1.46
[EBIT-TA]	0.18	0.15	1.21
[EBIT-Int]	-1.90	1.02	-1.87
[NI-TA]	**0.13**	**0.03**	**3.76**
[CA-CL-TA]	20.3	15.7	1.29

Table 5.1: The result of logit model

Statistics	value
Degree of freedom	1038
Variance	1068.4898
Log-Likelihood	-534.2449
Pearson	1457994.9823
R^2	0.2673
adj.R^2	0.2582
AIC	1096.4898
BIC	1165.9081
iterations	5
distinct obs.	1052

Table 5.2: The statistics of logit fit

5.2 Estimation Probit Model

Logit and probit models yield almost identical results and the choice of the model is usually arbitrary. Note that the parameters of the two models are scaled differently. Figure (5.2) shows the transformation functions of the probit and logit model.

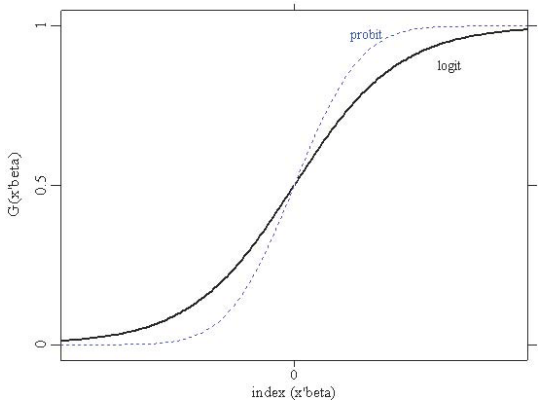

Figure 5.2: The transformation function in the probit and logit model

logit and probit1.xpl

From this plot we see that the curve of the probit model is similar to the curve of the logit model. The probit and logit models tend to produce very similar predictions. The parameter estimates in a logit model tend to be 1.6 to 1.8 times higher than they are in a corresponding probit model. Now for our data we estimate the parameters with probit model as in the table (5.3), in this table we see that the variable X_8 (total liabilities to total assets ratio) has a big influence on bankruptcy state of the the company and another variable X_1, X_2, X_4, X_6, X_{13} have a big influence. We can say, that we had obtained the same result in probit model.

Varible	Parameter (B)	t-value
[const]	-0.67	-1.07
[Cash-TA]	**2.87**	**2.86**
[Inv-TA]	**1.93**	**2.15**
[CA-TA]	-30.4	-2.56
[Kap-TA]	**-2.76**	**-2.59**
[Intg-TA]	0.14	0.32
[Log-TA]	**-0.06**	**-3.63**
[CL-TA]	7.83	0.64
[TL-TA]	**1.42**	**8.12**
[S-TA]	0,09	0.79
[EBIT-TA]	0.05	1.07
[EBIT-Int]	-2.80	-1.97
[NI-TA]	**0.04**	**2.25**
[CA-CL-TA]	8.11	0.87

Table 5.3: The result of probit model

At the end, probit and logit models are similar to one another, probit and logit are techniques for estimating the effects of a set of independent variables on a binary or dichotomous dependent variable. When OLS is used to estimate a binary dependent variable model, the model is often called a linear probability model (LPM). Probit and logit avoid several statistical problems with LPM and generally yield results that make more sense.

Chapter 6

Some Cases of Link

6.1 Computing GPLM Estimates

The Generalized Partial Linear Models $(GPLM)$ extends the (GLM) by a nonparametric component

$$P(Y \mid X, T) = G\{X^{\top}\beta + m(T)\}$$

Where $E(Y \mid X, T)$ denotes the expected value of the dependent variable given vectors of explanatory variables. The index $X^{\top}\beta + m(T)$ is linked to the dependent variable Y via a known function $G(.)$ which is called the link function in analogy to generalized linear models (GLM). There is in XploRe the gplm quantlib for estimating Generalized Partial Linear Models. We use bankruptcy data to illustrate the $GPLM$ estimation, and obtain the next plot in Figure (6.1). Table 6.1 shows the estimation results for $GPLM$. We consider that the variable X_6 is constant and consider t the variable X_1. From this table we say that the variables X_8 (total liabilities to total assets ratio) has an important influence on bankruptcy state of company and the variables X_2 (inventories to total assets ratio), X_4 (property, plant and equipment to total assets ratio), X_{12} (EBIT to Interest Payments ratio) and X_{13} (net income to total assets ratio) have a big influence too.

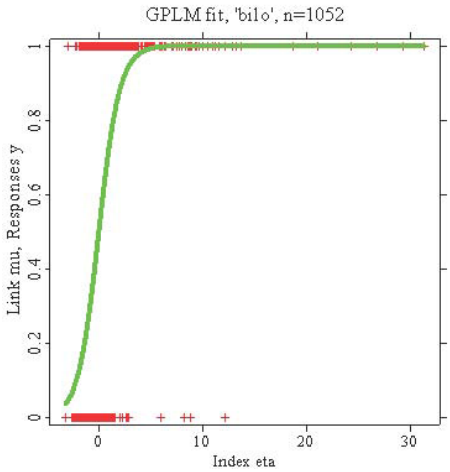

Figure 6.1: GPLM logit

GPLM.xpl

Estimates	B	s.e.	t-value
[Inv-TA]	**1.19**	**0.66**	**1.79**
[CA-TA]	**-18.9**	**15.5**	**-1.22**
[Kap-TA]	**-1.51**	**0.57**	**-2.65**
[Intg-TA]	0.08	0.71	0.12
[CL-TA]	15.5	15.5	1.00
[TL-TA]	**3.45**	**0.28**	**12.3**
[S-TA]	0.09	0.11	0.81
[EBIT-TA]	0.09	0.11	0.81
[EBIT-Int]	**-2.38**	**1.11**	**-2.14**
[NI-TA]	**0.12**	**0.04**	**3.24**
[CA-CL-TA]	17.6	15.5	1.13

Table 6.1: The result of GPLM model

Plots from $m(T)$ for $T = X_1$, $T = X_2$, $T = X_4$, $T = X_5$, $T = X_6$, $T = X_8$, $T = X_{10}$, $T = X_{11}$, $T = X_{12}$, $T = X_{13}$, are presented in figure 6.2 to figure 6.11 respectively. From these figures we say that the variables X_1, X_2, X_4 are almost linear, but the other variables are not linear.

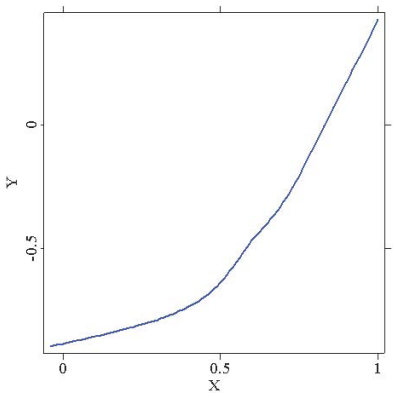

Figure 6.2: Plot from $m(T)$ for $T = X_1$

m1(t).xpl

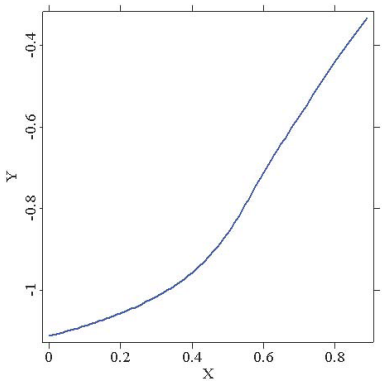

Figure 6.3: Plot from $m(T)$ for $T = X_2$

m1(t).xpl

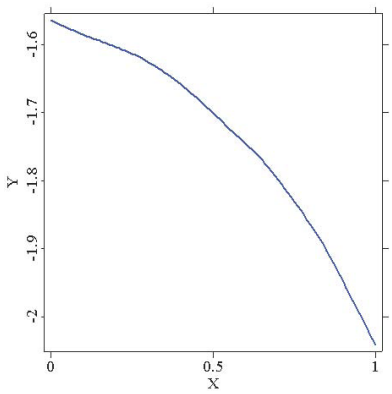

Figure 6.4: Plot from $m(T)$ for $T = X_4$

m1(t).xpl

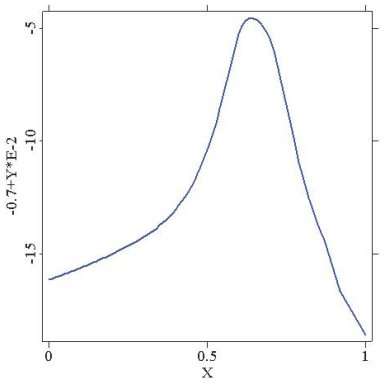

Figure 6.5: Plot from $m(T)$ for $T = X_5$

m1(t).xpl

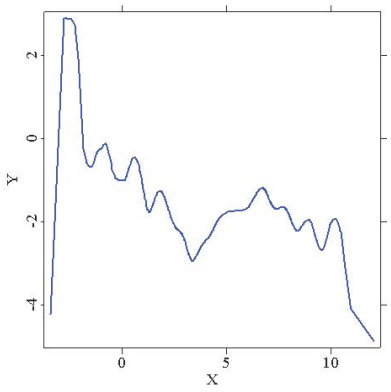

Figure 6.6: Plot from $m(T)$ for $T = X_6$

m1(t).xpl

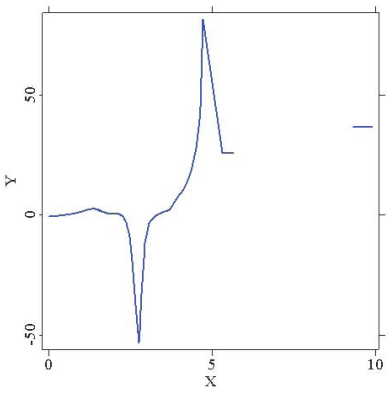

Figure 6.7: Plot from $m(T)$ for $T = X_8$

m1(t).xpl

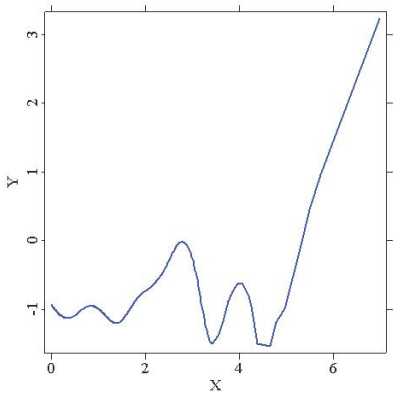

Figure 6.8: Plot from $m(T)$ for $T = X_{10}$

m1(t).xpl

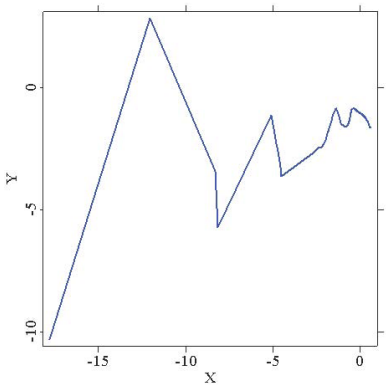

Figure 6.9: Plot from $m(T)$ for $T = X_{11}$

m1(t).xpl

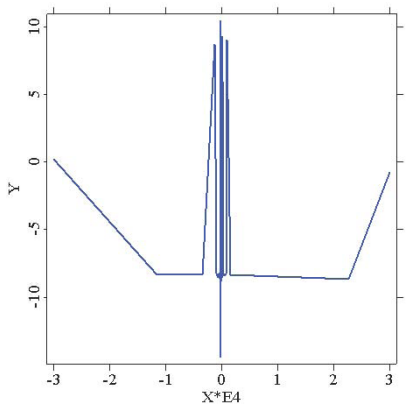

Figure 6.10: Plot from $m(T)$ for $T = X_{12}$

m1(t).xpl

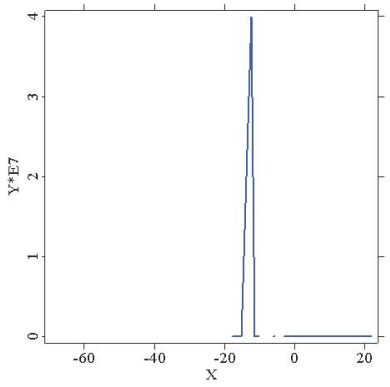

Figure 6.11: Plot from $m(T)$ for $T = X_{13}$

m1(t).xpl

6.2 Classification Rating

We present results of three classification ratings for GPLM model, logit model, and SVM. We use misclassification probabilities and the actual error rate (ARE), for the classification problem for the bankruptcy data. We present results for the GPLM model in table 6.2

		estimation	
		Bankrupt	Non-bankrupt
data	Bankrupt	305	121
	Non-bankrupt	190	436

Table 6.2: Classification rating for GPLM model

From this table we see that estimates of the misclassification probabilities are given by

$$\hat{p_{12}} = \frac{n_{12}}{n_2}$$

And

$$\hat{p_{21}} = \frac{n_{21}}{n_1}$$

And the actual error rate (AER) is given by

$$\frac{n_{12} + n_{21}}{n_2 + n_1} = \frac{121 + 190}{626 + 426} = 0.29$$

For the logit model we have the table (6.3) for the Classification rating, the ratio AER

		estimation	
		Bankrupt	Non-bankrupt
data	Bankrupt	310	116
	Non-bankrupt	157	469

Table 6.3: Classification rating for logit model

$$\frac{116 + 157}{626 + 426} = 0.25$$

And for SVM we have the table (6.4) for the Classification rating and the

		estimation	
		Bankrupt	Non-bankrupt
data	Bankrupt	317	109
	Non-bankrupt	140	486

Table 6.4: Classification rating for SVM

ratio AER

$$\frac{109 + 140}{626 + 426} = 0.23$$

We observe that classification ratings with SVM is superior to result presented for the logit model. This performance for SVM is as a result of the use of classifiers that provide high margin. However the logit model method gives a good alternative to SVM.

6.3 Conclusion

The logit model of bankruptcy prediction is a useful model to investors, analysts, and auditors. However, its results are only as accurate as the completeness of the data in the model. However, it should be noted that bankruptcy prediction is not a complete solution to risk measurement. It is just one of many tools that the analyst should consider in evaluating the effectiveness of management and the risk associated with an investment opportunity. This study has shown that profitability, leverage and Liquidity have important effects on the probability of bankruptcy. We had seen that SVM was a better method for classification rating, but on the other hand the logit model was a good method because we had got almost the same results.

6.4 References

Altman, E., *Financial Ratios, Discriminate Analysis and the Prediction of Corporate Bankruptcy*, The Journal of Finance 23 (September 1968): 589 - 609.

Beaver, W., (1966). *Financial ratios as predictors of failure. Empirical Research in Accounting* Selected Studies, 1966, supplement to vol.5, Journal of Accounting Research, pp. 71-111.

Christensen, R., (1990). *Log-Linear Models*, Springer-Verlag, New York.

Cook, Roy A. and Jeryl L. Nelson., *A Conspectus of Business Failure Forecasting*, 12 April 1998.

Härdle, W., Müller, M., Sperlich, S. and Werwatz, A. (2004). *Nonparametric and Semiparametric Models*, Springer-Verlag, Heidelberg.

Müller, M., (2000). *Semiparametric Extensions to Generalized Linear Models*.

Härdle, W., Hlavka, Z., Klinke, S. (2003). *Xplore Application Guide*, Springer-Verlag, Heidelberg.

Hart, O. (1999). *Different approaches to bankruptcy*, Annual World Bank Conference on Development Economics, Paris, june 21-23.

Kohler, U., (2002). *Ordinal Response Models*, Uni-Mannheim.

Lo, Andrew W. *Logit Versus Discriminant Analysis: A Specification Test*

and Application to Corporate Bankruptcies, Journal of Econometrics 31 (March 1986): 151 - 179.

Ohlson, J.,(1980). *Financial ratios and the probabilistic prediction of bankruptcy*, Journal of Accounting Research Spring, p.109-131.

Peaucelle, I.,(2005). *Dynamic analysis of bankruptcy and economic waves*, working paper N° 2005 - 09.

Platt, H., Platt, M., and Pedersen, J. June (1994). *Bankruptcy discrimination with real variables*.

Schmidheiny, K., (2004). *Binary Response Models*, Uni-De Lausanne, HEC-Applied Econometrics.

Sheppard, P., *The Dilemma of Matched Pairs and Diversified Firms in Bankruptcy Prediction Models*,The Mid-Atlantic Journal of Business 30 (March 1994): 9.

Stickney, Claude P., *Financial Reporting and Statement Analysis*. 3rd Edition. Ft. Worth, TX: The Dryden Press, 1996.